Previous Discussion

- Kleene Star Closure, Plus operation, recursive definition of languages, INTEGER, EVEN, factorial, PALINDROME, {$\mathbf{a^n b^n}$}, languages of strings (i) ending in a, (ii) beginning and ending in same letters, (iii) containing aa or bb (iv)containing exactly aa,

Explanation to Problems (Previous Lecture)

- Q)

1) Let S={ab, bb} and T={ab, bb, bbbb} Show that $S^* = T^*$ [Hint $S^* \subseteq T^*$ and $T^* \subseteq S^*$]

2) Let S={ab, bb} and T={ab, bb, bbb} Show that $S^* \neq T^*$ But $S^* \subset T^*$

Solution: Since $S \subset T$, so every string belonging to S^* , also belongs to T^* but bbb is a string belongs to T^* but does not belong to S^*.

Cont…

- 3) Let S={a, bb, bab, abaab} be a set of strings. Are abbabaabab and baabbbabbaabb in S^*? Does any word in S^* have odd number of b's?

  Solution: since abbabaabab can be grouped as (a)(bb)(abaab)ab , which shows that the last member of the group does not belong to S, so abbabaabab is not in S^*, while baabbbabbaabb can not be grouped as members of S, hence baabbbabbaabb is not in S^*. Since each string in S has even number of b's so there is no possiblity of any string with odd number of b's to be in S^*.

Cont…

Q1)Is there any case when S^+ contains Λ? If yes then justify your answer.

Solution: consider S={Λ,a} then

S^+ ={Λ, a, aa, aaa, …}

Here Λ is in S^+ as member of S. Thus Λ will be in S^+, in this case.

Cont…

Q2) Prove that for any set of strings S

i. $(S^+)^*=(S^*)^*$

Solution: In general Λ is not in S^+, while Λ does belong to S^*. Obviously Λ will now be in $(S^+)^*$, while $(S^*)^*$ and S^* generate the same set of strings. Hence $(S^+)^*=(S^*)^*$.

Q2) continued...

ii) $(S^+)^+ = S^+$

Solution: since S^+ generates all possible strings that can be obtained by concatenating the strings of S, so $(S^+)^+$ generates all possible strings that can be obtained by concatenating the strings of S^+, will not generate any new string.

Hence $(S^+)^+ = S^+$

Q2) continued…

iii) Is $(S^*)^+=(S^+)^*$

Solution: since Λ belongs to S^* ,so Λ will belong to $(S^*)^+$ as member of S^* .Moreover Λ may not belong to S^+, in general, while Λ will automatically belong to $(S^+)^*$.

Hence $(S^*)^+=(S^+)^*$

Regular Expression

- As discussed earlier that a^* generates
  Λ, a, aa, aaa, …
  and a^+ generates a, aa, aaa, aaaa, …, so the language $L_1 = \{\Lambda, a, aa, aaa, \ldots\}$ and $L_2 = \{a, aa, aaa, aaaa, \ldots\}$ can simply be expressed by a^* and a^+, respectively.
  a^* and a^+ are called the regular expressions (RE) for L_1 and L_2 respectively.
  Note: a^+, aa^* and a^*a generate L_2.

Recursive definition of Regular Expression(RE)

Step 1: Every letter of Σ including Λ is a regular expression.

Step 2: If r_1 and r2 are regular expressions then

1. (r_1)

2. $r_1 r_2$

3. $r_1 + r_2$ and

4. r_1^*

are also regular expressions.

Step 3: Nothing else is a regular expression.

Defining Languages (continued)...

- **Method 3 (Regular Expressions)**
  - Consider the language $L=\{\Lambda, x, xx, xxx, \ldots\}$ of strings, defined over $\Sigma = \{x\}$.

    We can write this language as the Kleene star closure of alphabet Σ or $L=\Sigma^*=\{x\}^*$

    this language can also be expressed by the regular expression x^*.
  - Similarly the language $L=\{x, xx, xxx, \ldots\}$, defined over $\Sigma = \{x\}$, can be expressed by the regular expression x^+.

- Now consider another language L, consisting of all possible strings, defined over $\Sigma = \{a, b\}$. This language can also be expressed by the regular expression

$$(a + b)^*.$$

- Now consider another language L, of strings having exactly double a, defined over $\Sigma = \{a, b\}$, then it's regular expression may be

$$b^*aab^*$$

Remark

- It may be noted that a language may be expressed by more than one regular expressions, while given a regular expression there exist a unique language generated by that regular expression.
- Example: all possible combination of {a,b}
- L1 = (a+b)*
- L2 = (a+b)* a* (a+b)* b*
- Both L1 and L2 are regular expressions for same language

Its Your Turn Now!

- language L, of even length, defined over $\Sigma = \{a, b\}$
- language L, of odd length, defined over $\Sigma = \{a, b\}$
- the language, defined over $\Sigma=\{a , b\}$ of words having at least one a
- the language, defined over $\Sigma = \{a, b\}$ of words having at least one a and one b
- the language, defined over $\Sigma=\{a, b\}$, of words starting with double a and ending in double b

Its Your Turn Now!

- the language, defined over Σ={a, b} of words starting with a and ending in b OR starting with b and ending in a
- Consider the language, defined over Σ={a, b} of words beginning with a
- the language, defined over Σ={a, b} of words beginning and ending in same letter
- the language, defined over Σ={a, b} of words ending in b
- the language, defined over Σ={a, b} of words not ending in a
- Language of strings, defined over Σ={a, b} having even number of a’s and even number of b’s

- Now consider another language L, of even length, defined over Σ = {a, b}, then it’s regular expression may be

$$((a+b)(a+b))^*$$

- Now consider another language L, of odd length, defined over Σ = {a, b}, then it’s regular expression may be

$$(a+b)((a+b)(a+b))^* \text{ or}$$

$$((a+b)(a+b))^*(a+b)$$

- Example:
  - Consider the language, defined over $\Sigma=\{a , b\}$ of words having at least one a, may be expressed by a regular expression $(a+b)^*a(a+b)^*$.
  - Consider the language, defined over $\Sigma = \{a, b\}$ of words having at least one a and one b, may be expressed by a regular expression

$(a+b)^*a(a+b)^*b(a+b)^*+ (a+b)^*b(a+b)^*a(a+b)^*$.

- Consider the language, defined over $\Sigma=\{a, b\}$, of words starting with double a and ending in double b then its regular expression may be $aa(a+b)^{*}bb$
- Consider the language, defined over $\Sigma=\{a, b\}$ of words starting with a and ending in b OR starting with b and ending in a, then its regular expression may be $a(a+b)^{*}b+b(a+b)^{*}a$

TASK

- Consider the language, defined over $\Sigma=\{a, b\}$ of **words beginning with a**, then its regular expression may be $a(a+b)^*$

- Consider the language, defined over $\Sigma=\{a, b\}$ of **words beginning and ending in same letter**, then its regular expression may be $(a+b)+a(a+b)^*a+b(a+b)^*b$

TASK

- Consider the language, defined over $\Sigma=\{a, b\}$ of **words ending in b**, then its regular expression may be $(a+b)^*b$.
- Consider the language, defined over $\Sigma=\{a, b\}$ of **words not ending in a**, then its regular expression may be $(a+b)^*b + \Lambda$. It is to be noted that this language may also be expressed by $((a+b)^*b)^*$.

An important example

The Language EVEN-EVEN :

Language of strings, defined over Σ={a, b} having **even number of a's and even number of b's**. *i.e.*

EVEN-EVEN = {Λ, aa, bb, aaaa,aabb,abab, abba, baab, baba, bbaa, bbbb,…} ,

its regular expression can be written as

$(aa+bb+(ab+ba)(aa+bb)^*(ab+ba))^*$

Note

- It is important to be clear about the difference of the following regular expressions

$r_1=a^*+b^*$

$r_2=(a+b)^*$

Here r_1 does not generate any string of concatenation of a and b, while r_2 generates such strings.

Equivalent Regular Expressions

- **Definition:**

  Two regular expressions are said to be equivalent if they generate the same language.

  Example:

  Consider the following regular expressions

  $r_1 = (a + b)^* (aa + bb)$

  $r_2 = (a + b)^* aa + (a + b)^* bb$ then

  both regular expressions define the language of strings **ending in aa or bb**.

Note

- If $r_1 = (aa + bb)$ and $r_2 = (a + b)$ then

1. $r_1 + r_2 = (aa + bb) + (a + b)$
2. $r_1 r_2 = (aa + bb)\ (a + b)$
   $= (aaa + aab + bba + bbb)$
3. $(r_1)^* = (aa + bb)^*$

Regular Languages

- **Definition**:

  The language generated by any regular expression is called a **regular language**.

  It is to be noted that if r_1, r_2 are regular expressions, corresponding to the languages L_1 and L_2 then the languages generated by $r_1 + r_2$, r_1r_2(or r_2r_1) and r_1^*(or r_2^*) are also regular languages.

Note

- It is to be noted that if L_1 and L_2 are expressed by $r_1$and r_2, respectively then the language expressed by
  1) $r_1 + r_2$, is the language $L_1 + L_2$ or $L_1 \cup L_2$
  2) $r_1 r_2$, , is the language $L_1 L_2$, of strings obtained by prefixing every string of L_1 with every string of L_2
  3) r_1^*, is the language L_1^*, of strings obtained by concatenating the strings of L, including the null string.

Example

- If r_1=(aa+bb) and r_2=(a+b) then the language of strings generated by r_1+r_2, is also a regular language, expressed by (aa+bb)+(a+b)
- If r_1=(aa+bb) and r_2=(a+b) then the language of strings generated by r_1r_2, is also a regular language, expressed by (aa+bb)(a+b)
- If r=(aa+bb) then the language of strings generated by r^*, is also a regular language, expressed by $(aa+bb)^*$

All finite languages are regular.

Example:

Consider the language L, defined over Σ={a,b}, of strings of length 2, **starting with a**, then

L={aa, ab}, may be expressed by the regular expression aa+ab. Hence L, by definition, is a regular language.

Note

It may be noted that if a language contains even thousand words, its RE may be expressed, placing ‘ + ’ between all the words.

Here the special structure of RE is not important.

Consider the language L={aaa, aab, aba, abb, baa, bab, bba, bbb}, that may be expressed by a RE aaa+aab+aba+abb+baa+bab+bba+bbb, which is equivalent to (a+b)(a+b)(a+b).